

Automatic Breast Lesion Evaluation for Comparative Studies

BACHELORARBEIT

zur Erlangung des akademischen Grades

Bachelor of Science

im Rahmen des Studiums

Medizinische Informatik

eingereicht von

Thomas Tramberger

Matrikelnummer 01326870

an der Fakultät für Informatik
der Technischen Universität Wien

Betreuung: Univ.Ass. Renata Raidou, MSc PhD

Wien, 8. Mai 2018

Thomas Tramberger

Renata Raidou

Automatic Breast Lesion Evaluation for Comparative Studies

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Medical Informatics

by

Thomas Tramberger

Registration Number 01326870

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Ass. Renata Raidou, MSc PhD

Vienna, 8th May, 2018

Thomas Tramberger

Renata Raidou

Erklärung zur Verfassung der Arbeit

Thomas Tramberger
Voralpensiedlung 16, 3350 Haag, Austria

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 8. Mai 2018

Thomas Tramberger

Danksagung

Ich bedanke mich herzlich bei meiner Familie, die mir mein Studium ermöglichte und mich immer unterstützt hat, bei meiner Freundin für die vielen motivierenden Worte und bei meiner Betreuerin Renata Raidou für die hervorragende Unterstützung.

Acknowledgements

I wish to express my sincere thanks to my family for the opportunity to continue my education, my girlfriend for all the encouraging words and my supervisor Renata Raidou for her outstanding support.

Kurzfassung

Brustkrebs ist die am häufigsten auftretende Krebsart mit einer hohen Mortalitätsrate. Neoadjuvante Chemotherapie wird vor der operativen Entfernung des Tumors angewendet um die Größe des Tumors zu verringern. Derzeit finden viele Studien statt, die sich mit unterschiedlichen Chemotherapie Strategien beschäftigen. In dieser Arbeit wurde eine Software entwickelt welcher der Analyse und dem Vergleich der Auswirkungen dieser Behandlungen dient. Die Studiendaten liegen anhand von 4D Dynamic Contrast-Enhanced Magnetic Resonance Imaging Daten vor. Um den Arbeitsaufwand der Segmentierung und Verknüpfung der Läsionen über die Zeit zu verringern wurde ein automatisches Verfahren implementiert. Dieses nutzt das Prinzip der time-signal intensity curve und morphologische Features zur Classification von Läsionen durch eine Support Vector Machine. Zur Analyse der Ergebnisse dienen zwei Sichten. Die Intra-patient View dient der Analyse des Tumorverhaltens einzelner Patienten über die Zeit. Die Multi-patient View bietet den Vergleich mehrerer Patienten und inkludiert zusätzlich importierte Patientendaten. Beide Sichten sind mittels JavaScript jederzeit erweiterbar. Aufgrund fehlender Segmentierungen durch einen Pathologen konnte keine Evaluierung der automatischen Segmentierungsmethode erfolgen.

Abstract

Breast cancer is the most common cancer with a high mortality rate. Neoadjuvant chemotherapy is conducted before surgery to reduce the breast tumor mass. Currently, a lot of trials are taking place, with the purpose of understanding the effects of different chemotherapy strategies. In this work a software is developed to analyse and compare the influence of these treatments. The study data is available as 4D Dynamic Contrast-Enhanced Magnetic Resonance Imaging data. To reduce the time of manual segmentation and the connection of segmented lesions over time a automatic procedure was implemented. This process uses the time-signal intensity curve and a support vector machine to classify lesions with calculated morphological features. To analyse the data, two views are available. The Intra-patient view visualizes the tumor behaviour of an individual patient over time. With the Multi-patient view the user is able to compare multiple patients' lesions and additional added patient data. Both views are implemented with JavaScript and can be expanded easily. Because of missing ground truth an evaluation of the automatic segmentation method was not possible.

Contents

Kurzfassung	xi
Abstract	xiii
Contents	xv
1 Introduction	1
1.1 Motivation	1
1.2 Aim of the work and problem statement	2
1.3 Methodological approach	2
1.4 Data	3
2 State of the art	5
2.1 Breast segmentation	5
2.2 Lesion detection/segmentation	5
2.3 Data visualization	6
3 Methodology	7
3.1 Automatic Lesion Segmentation	7
3.2 Data visualization	13
4 Implementation	17
4.1 MeVisLab	17
4.2 Structure and Details	18
4.3 Limitations	19
5 Conclusion and future work	21
5.1 Conclusion and future work	21
List of Figures	23
Bibliography	25

Introduction

1.1 Motivation

With 5.480 new cases of breast cancer in 2015, being 30% of all cancer incidences, it is the most common cancer in Austria. Its high mortality rate cost 1.568 women and 22 men their lives in the year 2015 [Kre17]. At the end of 2015, 74.170 women and 647 men were diagnosed with breast cancer [Bru18].

Neoadjuvant therapy (NAT) is conducted before surgery to reduce the breast tumor mass, especially if the tumor is inoperable. Currently, a lot of trials are taking place, with the purpose of understanding the effect of different chemotherapy strategies on different patients and the response of the latter to the selected therapy method, with the course of time [LAA⁺14, LTR⁺08].

Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) is a common acquisition modality for breast cancer detection and monitoring of malignancy. Many papers address the topic of image segmentation to automatically detect and locate breast lesions. However, this work

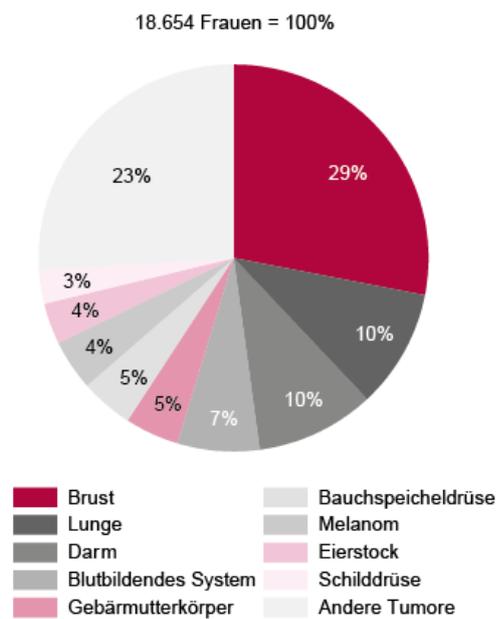


Figure 1.1: Cancer statistic, Austria 2015. Data from [Bru18]

aims at designing and implementing a visualization approach which enables the identification and analysis of breast lesions. The aim of this thesis is to develop a tool which allows the user to compare across a high-number of patients (multi-patient analysis) and to assess the response of the same patient through time (intra-patient follow-up study). These two topics have been rarely investigated in the past.

1.2 Aim of the work and problem statement

This work aims to develop a tool that provides a visual comparison among patients' cancer stages using automatic segmentation. This tool should help medical researchers to identify patterns when evaluating the applied therapies. Therefore, visualizations are needed for intra-patient analysis, to give an overview of a single patients lesions development, and for multi-patient analysis, to compare patients lesions.

To achieve this purpose the following challenges have to be tackled:

- Designing and implementing an efficient automatic segmentation workflow for correct detection and segmentation of multiple lesions.
- Treatment follow-up of segmented lesions of a patient over time.
- Because of imperfection of all segmentation approaches, the possibility of manual segmentation has to be given.
- Identification of appropriate visualization approaches to provide functionality for finding correlations between therapies, lesion changes and additional health record data of patients.

1.3 Methodological approach

The methodology of the work is structured as followed:

1. Literature review:

- Different approaches for medical image segmentation are studied for the implementation of the automatic segmentation of the breast lesions.
- State-of-the-art reports of comparative visualization are reviewed and used in the design of the proposed visualizations.

2. Planning and implementation:

The following tasks haven been executed in the following order:

- Development of a pipeline to automatically segment multiple breast lesions. The pipeline additionally connects segmented lesions over time, using given DCE-MRI data of the patients.

- Investigation for visualization methods to compare multiple patients lesion behaviour with additional health record data.
- Testing of the segmentation pipeline and the visualization techniques.

These tasks were implemented as a extension for the open-source software MeVisLab [MeVb].

1.4 Data

The data used in this study was provided by the cancer image archive TCIA [TCI17]. It includes longitudinal DCE-MRI data of 68 patients in treatment with neoadjuvant chemotherapy (NACT) for invasive breast cancer. As shown in the Figure 1.2, all patients received 4 circles of adriamycin-cytoxan administered every three weeks. A subset of 17 patients were additionally medicated with taxane afterwards before surgery. Overall, three DCE-MRI scan series for all patients as well as an extra scan were made for the taxane threaded subset of patients before surgery. The first scan happened before any treatment, the second scan after one cycle of chemotherapy and the third DCE-MRI scan after chemotherapy. Figure 1.2 gives an overview of the structure of the given data.

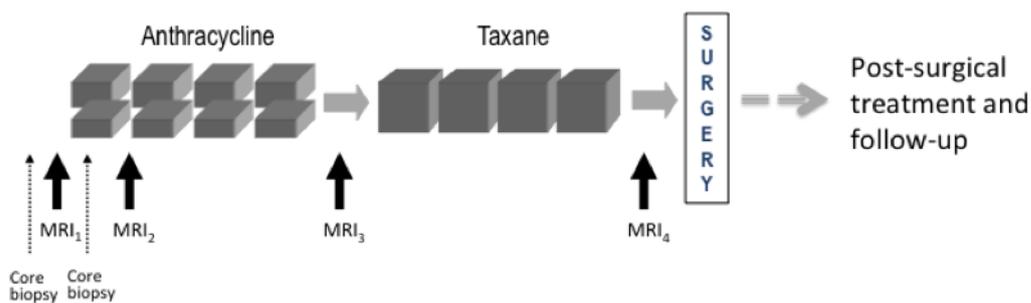


Figure 1.2: Schematic depiction of acquisition procedure. [TCI17]

For data acquisition a 1.5-T scanner was used, outputting T-1 weighted, fat-suppressed 3D sequences with an acquisition matrix of $256 \times 192 \times 60$, a slice thickness of 2 mm and a spatial resolution of $0.7 \times 0.94 \times 2.0 \text{ mm}^3$. Each scan series includes a total of three time points: a pre-contrast scan, followed by an early post-contrast scan after 2.5 minutes of inducing the contrast agent and a late post-contrast scan 7.5 minutes after induction [TCI17].

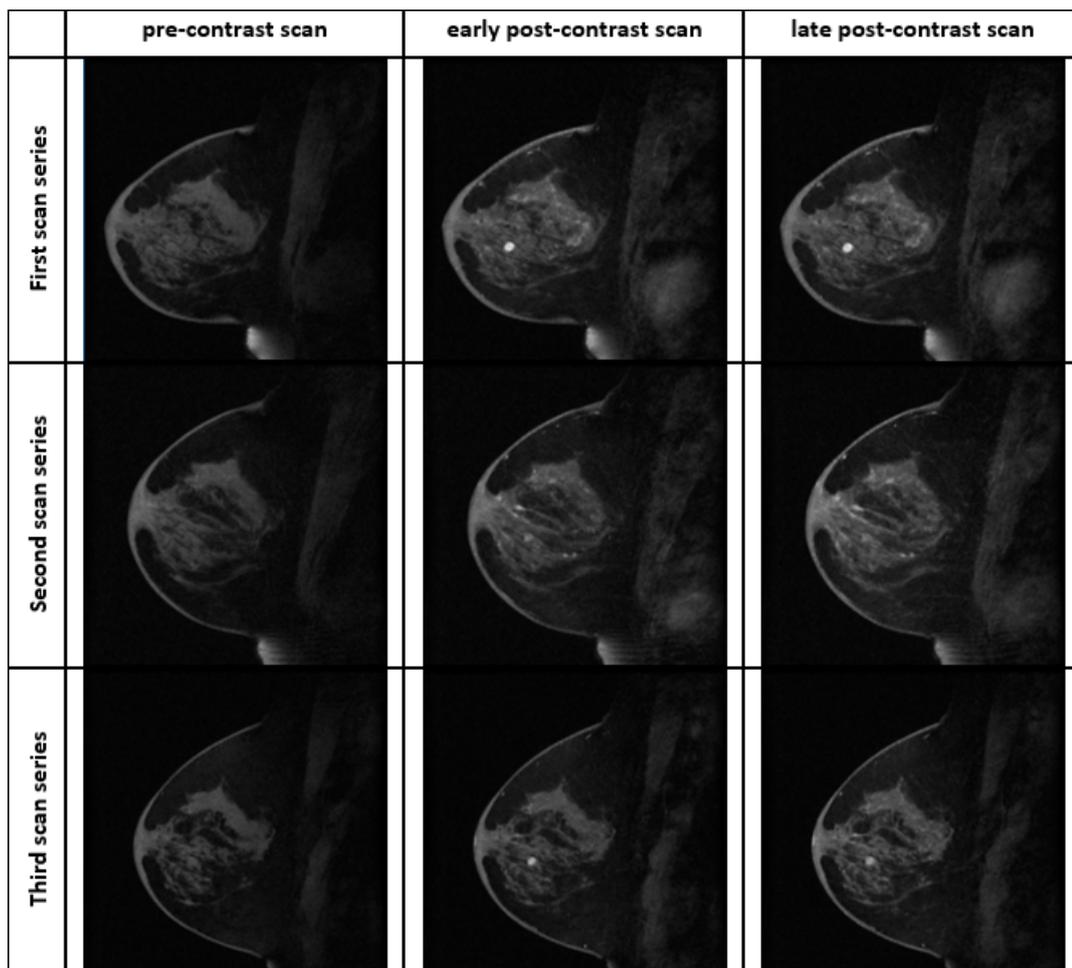


Figure 1.3: Example of the scanned DCE-MRI data of a patient

State of the art

The following chapter gives an overview on current research of related topics. It is divided into three main parts: The first part contains approaches, aiming to segment the breast from the body. Then several papers dealing with lesion detection and segmentation are presented. The last chapter discusses the topic of data visualization.

2.1 Breast segmentation

A wide range of proposed approaches for automatic breast segmentation in MRI exists, starting from simple methods to extensive tasks. A very simple strategy is the histogram-based segmentation by searching the coronal plane for the highest cumulative enhancement to find the heart and finally drawing a line along the superior-inferior axis to separate the chest from the body [WGOM15]. Although, histogram-based approaches are simple, they cannot be used on fat-saturated MRI data [VGD⁺11]. Further simple approaches are atlas-based segmentation [VGD⁺11] and using landmarks [TCS⁺18] e.g. the sternum for segmentation. Atlases, created by manual segmentation of the breast, heart, chest wall and lungs are applied based on the breast size [VGD⁺11, GMKM⁺15]. More complex papers using fuzzy C-means algorithms or hessian-based sheetness filters, with additional procedures, to segment the pectoral muscle boundaries or capturing the breast tissue [NCC⁺08, WPI⁺]. Also, papers aim to detect the chest wall line by utilizing edge-detection filters to divide the breast from the body [WWC⁺13].

2.2 Lesion detection/segmentation

As well as in breast segmentation, there is a great variety of automatic lesion detection and segmentation approaches. Most commonly used scientific foundation is the time-signal intensity curve (TIC) (see 3.1.2) combined with further processing for lesion detection. To screen the time-intensity-curve for lesions, mainly trained classifiers or

artificial neural networks are used [VGB⁺09, VGD⁺11, ZBE⁺07, JCJG14]. The work of Fusco et al. [FSF⁺16] focuses on pattern recognition approaches for lesion segmentation by systematically reviewing related papers. It gives an overview about the different used classifiers, the resulted sensitivity and specificity. The results show that linear discriminant analysis (LDA) and tree-based (TC) perform best, moreover simultaneous use of dynamic and morphological features achieve best accomplishments. Even though, these approaches lead to the best results, each type of used classifier, covering support vector machine (SVM), artificial neural networks (ANN) and Bayesian classifiers (BC), and feature types were able to reach a sensitivity over 90% and a specificity over 88%.

2.3 Data visualization

Many papers deal with lesion detection and segmentation, skipping the visual comparison of the results and the representation of perfusion data e.g. DCE-MRI data and health record information. However, Preim et al. [POM⁺09] introduces existing methods for CT and MR perfusion data analysis and visualization. Basic visualization techniques are cine movies, subtraction images and color-coded parameter maps. Furthermore, a glyph-based approach is featured in the work of Glaßer et al. [GSO⁺09] to display voxel-wise relative enhancement. Theus [The81] covers the issue of high dimensional data visualization, proposing the parallel coordinate plots. Van Wijk and Van Selow [VV] address the representation of time series data, presenting calendar based analysis. For health record data, Andry et al. [ANN⁺09] recommend the use of new web technologies for appealing, precise, but simple interactive demonstration.

Methodology

The following chapter describes the automatic breast cancer detection and segmentation of patients' MRI series. It gives in-depth explanations of the individual pipeline steps. In the subsection data visualization, the information visualization approach for the collected lesion data for each patient is explained and how additional added health record data is presented.

3.1 Automatic Lesion Segmentation

Figure 3.1 displays the developed breast lesion segmentation pipeline. As shown, the tumor detection is applied at the first series. Series refer to the collected DCE-MRI scan series of the patients (see 1.4). The first series of the DCE-MRI data undergoes multiple filtering processes, including breast segmentation, TIC-filtering and additional procedures. Afterwards, morphological features of the remaining areas are calculated and a classifier decides if the regions are cancerous. Finally, the found lesions are mapped to the subsequent transformed and filtered DCE-MRI images, searching for the identical lesion and connecting them over time.

3.1.1 Breast segmentation

Anatomical segmentation of the breast discards structures located outside of the region of interest (ROI). This includes anatomical structures, e.g. the heart, with similar TIC, due to high intensity enhancement (see Figure 3.2) [VGB⁺09]. By eliminating those areas, the speed of the process as well as the reliability of the classification is increased.

As shown in Figure 3.3, the idea of the breast segmentation pipeline is to segment the outline of the breast and simultaneously reduce the breast tissue. The first step is to use a sobel filter [SF68], followed by a morphological gradient operation [RSB92] and a threshold to outline the outer breast border, to reduce noise and the inner breast tissue.

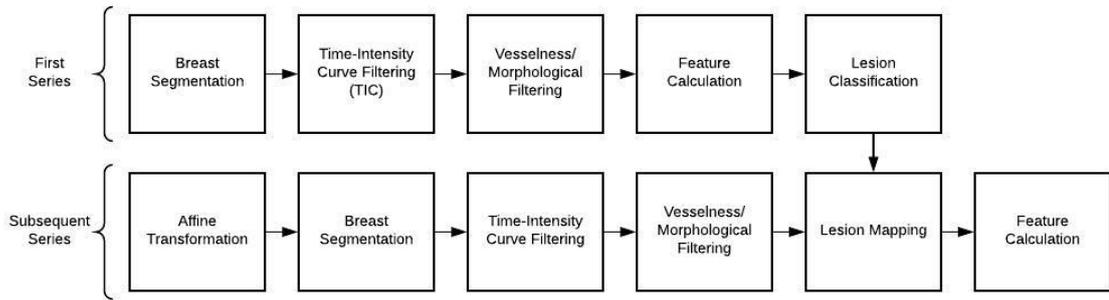


Figure 3.1: Pipeline of the automatic segmentation mechanism for an individual patient

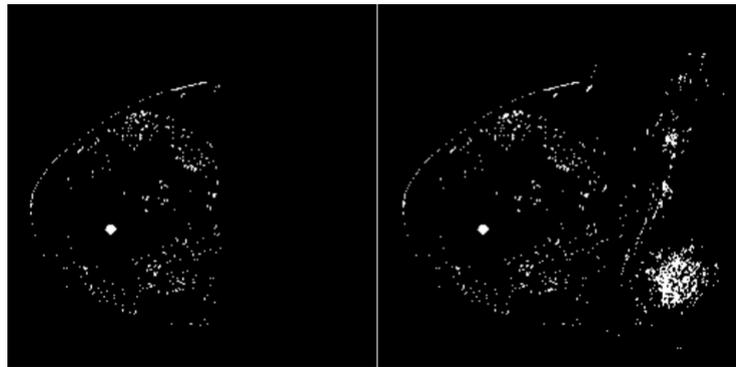


Figure 3.2: Comparison of TIC and vesselness filtered images with (left) and without (right) breast segmentation

Morphological filters, erosion and dilation, decrease the breast tissue even more. This also disconnects the inner breast tissue from the chest. Finally, the largest area, the outer breast with some remaining inner tissue, is selected and the minimum and maximum of each axes of the middle slice defines the ROI. Using the extremes of the middle slice instead for each slice individually eliminates inconsistencies especially of the outer breast slices caused by the DCE-MRI data or the previous process, trading accuracy for stability of the algorithm. Further, the speed is slightly increases.

The sobel filter parameters for minimum value and maximum value were set to 40 and 300. Dilation size and erosion size for the morphological gradient were both set to three. The threshold's value was set to 300, erosion and dilation filter both used a $2 \times 2 \times 1$ kernel. These parameters highly depend on the DCE-MRI data and are optimized for the used images. Changing the parameters for erosion and dilation could may lead to not separate the breast tissue from the chest.

3.1.2 Time-signal intensity curve

The TIC of DCE-MRI data displays the change of intensity caused by the injected contrast-enhancing agent. This shows large physiological variations, particularly depending on

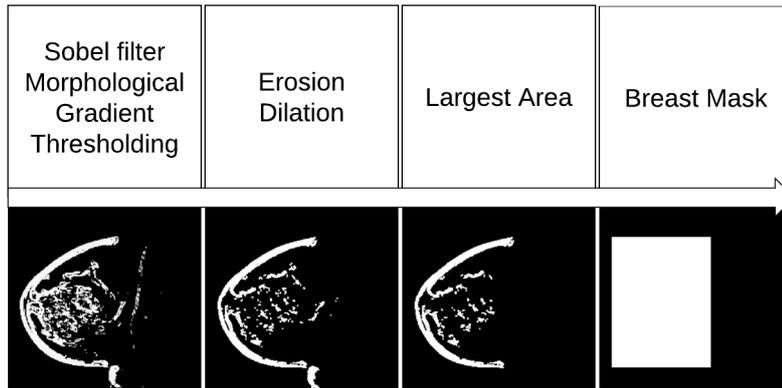


Figure 3.3: Breast segmentation pipeline

different vascular permeability [VGB⁺09].

The quality of the TIC is highly depending on the acquired data. A high temporal resolution results in a smoother TIC. Furthermore, when measuring a TIC of DCE-MRI data, it is important that the patient is moving as little as possible while acquiring a series (<3 pixels) [CL13]. This issue is not tackled in this work and is a requirement for chosen data.

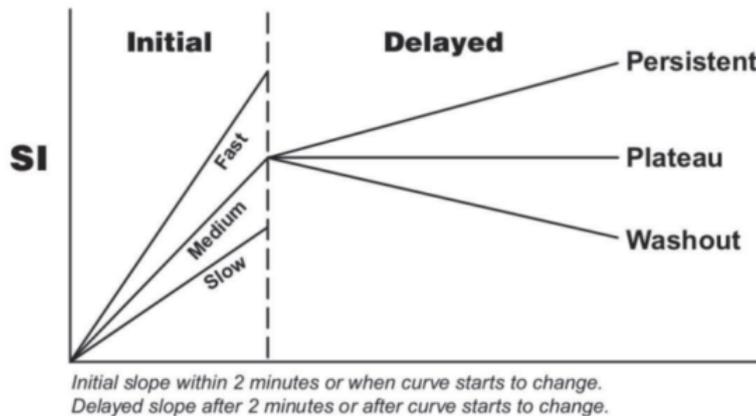


Figure 3.4: Visualization of the main phases of the TIC [CL13]

The interpretation of the TIC is not standardized. Essentially there are two main phases in the observation of the kinetic curve. The initial slope (2 minutes after agent injection) and the delayed phase as shown in Figure 3.4. The initial slope is separated into "slow", "medium" and "rapid". It indicates the issue of tumor angiogenesis. Rapid increase of intensity, e.g. 90%, highly suggest malign tissue. The delayed phase is divided into "persistent" (type 1), "plateau" (type 2) and "washout" (type 3). It reflects the formation of stromal tumor cells. Intensity behaviour of type 1 suggests benign tissue, while a type

3 behaviour denotes malign cells. A plateau curve indicates either one of them. Equation 3.1 and equation 3.2 display the formula for calculating the initial slope and the delayed slope, where SI_{t_0} is the intensity at the time of the contrast agent injection, SI_{t_1} the intensity after 2.5 minutes at the initial peak and $SI_{t_{last}}$ the intensity at the last recorded time point [CL13].

$$early\ enhancement = [(SI_{t_1} - SI_{t_0})/SI_{t_0}] * 100(\%) \quad (3.1)$$

$$delayed\ slope = [(SI_{t_{last}} - SI_{t_1})/SI_{t_1}] * 100(\%) \quad (3.2)$$

In this work the TIC was utilized to filter the image data and select only qualified voxels for further processing. The threshold for the initial peak was set to 70% and voxels of type 1 behaviour with delayed slope over 5%, as suggested by El Khouli et al. [EMJ⁺09], were removed.

3.1.3 Filtering

Tumors often develop new, disorganized blood supply with thinner walls which will appear in the TIC-filtered image. Therefore, a hessian vesselness filter [FNVV98] was executed on the processed image. Further a erosion filter, followed by a dilation filter, both with an external kernel of 2x2x2x1, were used to eliminate noise. Finally, small areas under the size of 40 voxels ($\sim 50\text{ mm}^3 \approx 5\text{ mm}$ diameter) were eliminated, adapted from the work of Vignati et al. [VGB⁺09, VGD⁺11].

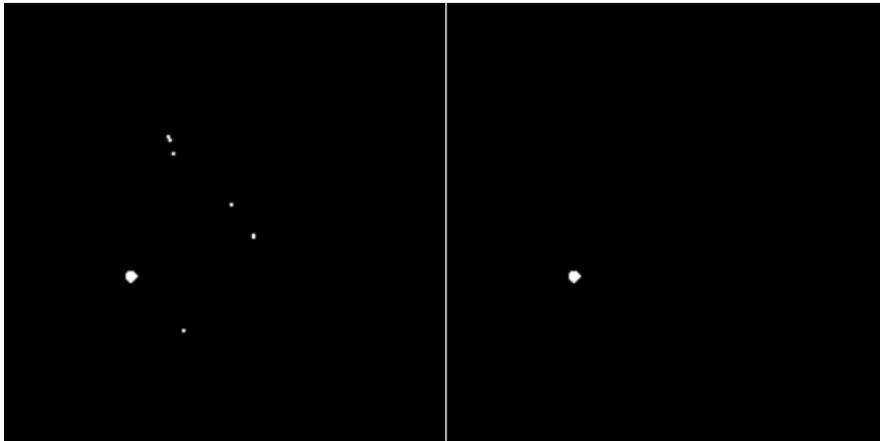


Figure 3.5: Image after morphological filters (left) and fully filtered image (right).

3.1.4 Classification

A SVM separates the remaining areas of the filtered DCE-MRI image by assigning them into two classes, cancerous and non-cancerous, by using morphological features presented

in section 3.1.5. The classification function is trained by labelled training data. A SVM is a classifier, initially created for binary problems, separating n-dimensional vector data [Yu09].

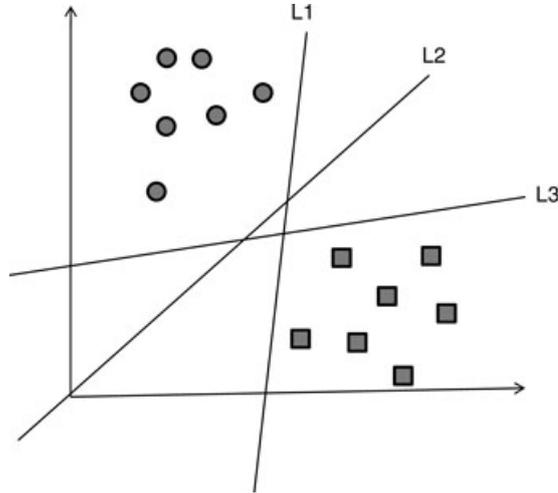


Figure 3.6: Linear, two-dimensional SVM.

The choice of using a SVM was supported by [FSF⁺16]. The paper compares several pattern recognition approaches of breast cancer classification, concluding that all selected classifier, including SVMs, can lead to reasonable results, even though LDA and TC yield to a more precise outcome. Moreover, the need to design a appropriate structure, e.g. when using an artificial neural networks, is not given. Furthermore, Juntu et al. [JSD⁺10] results show that the SVM out performances a neural network and a C4.5 decision tree, disagreeing with the conclusion of Fusco et al. [FSF⁺16].

3.1.5 Feature extraction

A total of six shape-based morphological features were calculated including Volume, Surface Area, Compactness, NRL (Normalized Radial Length) mean, Sphericity, NRL ratio. Those features originate from the work of [NCC⁺08] and represent 3D properties as well as additional geometric characteristics. The calculation of these features can be seen in the equations 3.3 - 3.7, with p defined as the pixel size on the image plane, t as the slice thickness and r as the individual radial length. F_{ROI} is the pixel number of the ROI, S_{ROI} is the pixel number along the boundary of the ROI and σ_{NRL} is the standard deviation of NRL. According to Fusco et al. [FSF⁺16], dynamic and morphological features combined achieve the best performance. Dynamical features are directly obtained from the TIC, e.g. maximum intensity ratio.

$$\text{Volume} \quad Vol = \sum_{\forall x, \forall y, \forall z} F_{ROI}(x, y, z) * p^2 * t \quad (3.3)$$

$$\text{Surface} \quad Surf = \sum_{\forall x, \forall y, \forall z} S_{ROI}(x, y, z) * p * t \quad (3.4)$$

$$\text{Compactness} \quad Comp = \frac{Surf}{Vol} \quad (3.5)$$

$$\text{NRL mean} \quad \mu_{NRL} = \frac{1}{N} \sum_{j=1 \dots N} r_j \quad (3.5)$$

$$\text{Sphericity} \quad Spher = \frac{\mu_{NRL}}{\sigma_{NRL}} \quad (3.6)$$

$$\text{NRL ratio} \quad R_{NRL} = \frac{1}{N * \mu_{NRL}} \sum_{j=1 \dots N} (r_j - \mu_{NRL}) : r_j > \mu_{NRL} \quad (3.7)$$

3.1.6 Subsequent Series

After lesion extraction of the patients' first DCE-MRI series, the found lesions are mapped to the following series. Before mapping the segmented lesions to the subsequent DCE-MRI image a 3D affine transformation and additional processing is applied (see Figure 3.1). Then, the area of the found tumors of the first series is mapped to the filtered image of a subsequent series, looking for nearby areas as presented in Figure 3.8. Therefore, all voxel coordinates from of the tumors bounding box are converted to the transformed image, searching for the lesion in the subsequent series.

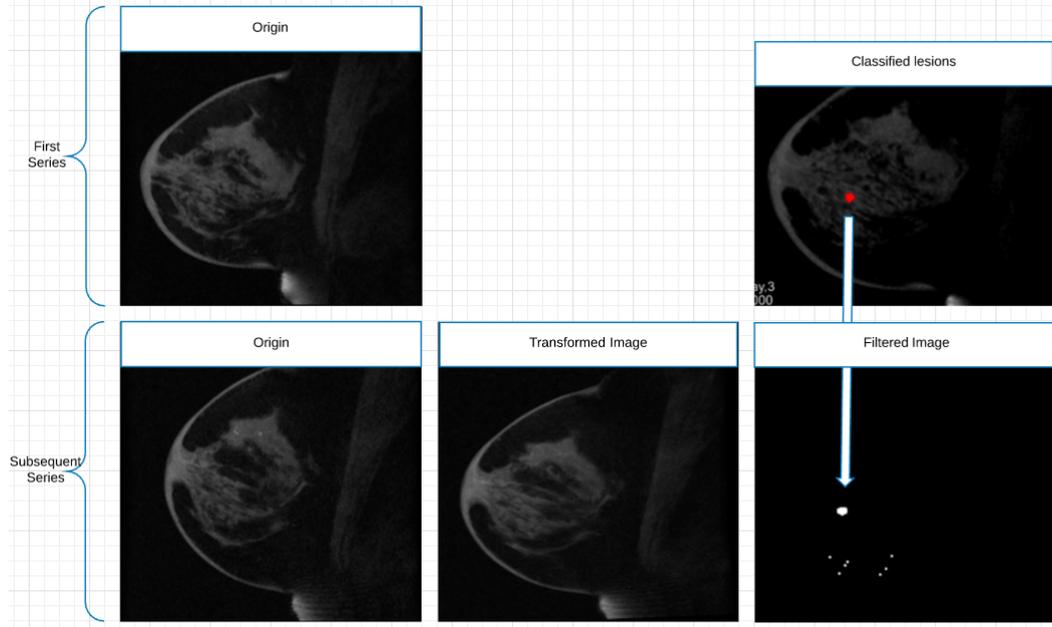


Figure 3.7: Lesion mapping to subsequent time

3.2 Data visualization

Visualization methods translate logical and complex data into a lucid format. They support us in the understanding of information in a more comprehensible way and helps in the finding of correlations. In this work the visualization is divided into an Intra-patient view and a Multi-patient view. The Intra-patient view provides charts indicating changes of patients' lesions.

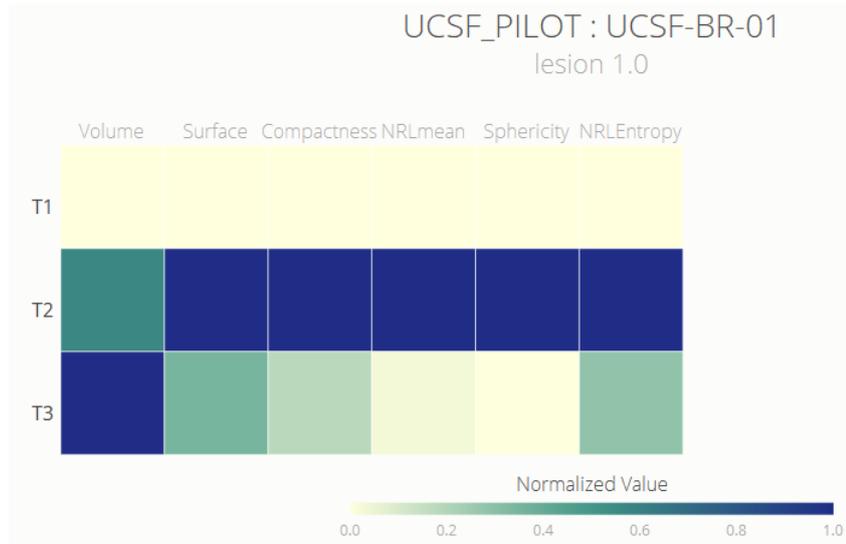


Figure 3.8: Calendar view

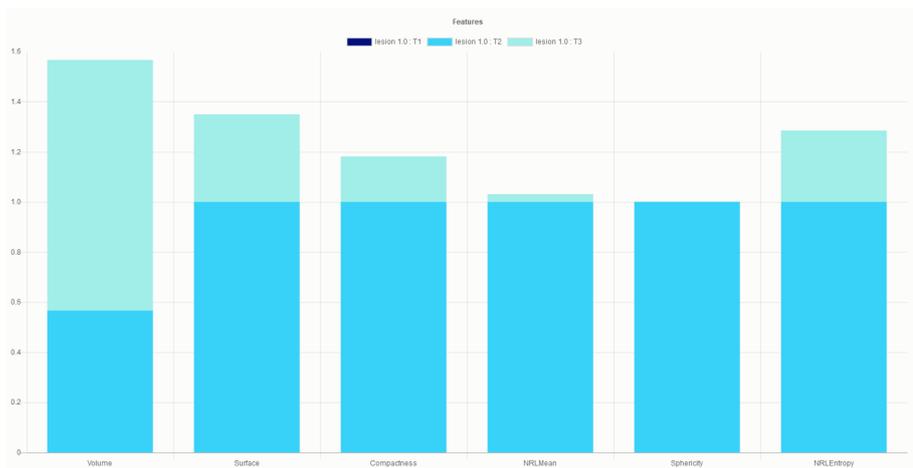


Figure 3.9: Bar plot

3.2.1 Intra-patient view

The feature data is normalized to eliminate the effects of certain gross influences and to display the degree of change more easily. The calendar view is a recommended method to analyze time series data and cluster analysis [VV]. Here it is used to help rating the applied chemotherapy for each patient. Rows represent the different morphological features (see 3.1.5), columns display the different series. Colors indicate the dimension of morphological feature change, making it simple to identify those changes. For each lesion of a patient a calendar view is created.

The grouped, stacked bar plot displays the change at each time point, also allow to compare multiple lesions of a patient. The encoding of quantities by length is a highly accurate representation. [SG14] Therefore, compared to the calendar view, the exact percentage of lesion growth/reduction and the precise rate of feature changes can be withdrawn by the user. The group property supports the comparison of several lesions at once. In the implemented bar plot the color of the bars indicate the rate of change from series to series. Dark colors visualize early changes, lighter colors later changes of features. Bars of the same features, but from different lesions are grouped to put them next to each other for a improved comparative view. The axes display the morphological features as well as the percentage of change.

3.2.2 Multi-patient view

The Multi-patient view gives the opportunity to compare data of multiple patients, also not lesion-related data e.g. the patients race, mainly to give the potential to discover significant correlations not only in lesion behaviour. An example of an appropriate plot assisting to find correlations is shown in Figure 3.10. Even though, it is traditionally used to visualize flow of energy or material flow, the Sankey diagram illustrates quantitative information about the relationship of patient's attributes, highlighting the distribution and connections between [RHF, CFR⁺13, Sch08]. In this work a interactive circular Sankey diagram (Figure 3.10) was implemented. The labelled outer cycles represent selected attributes, the various forms of an attribute are shown in the same color. The user can interact by hovering over an attribute to display the connections between two options of different attributes. Figure 3.11 gives an example of an use case of the interactive use of the Sankey diagram.

Another approach features the PCP. The method of PCPs is well-known for high-dimensional data [The81, Ins85]. The patients can be selected individually. Their lesions are then presented in the PCP displaying the values of the morphological features on the x-axes, enabling to find correlations or trends to grade the applied therapy. Red lines show data from the first, blue from the second, green from the third and orange from the forth series. The abilities of move x-axes and to restrict data by interval of feature values are given (Figure 3.13).

3. METHODOLOGY

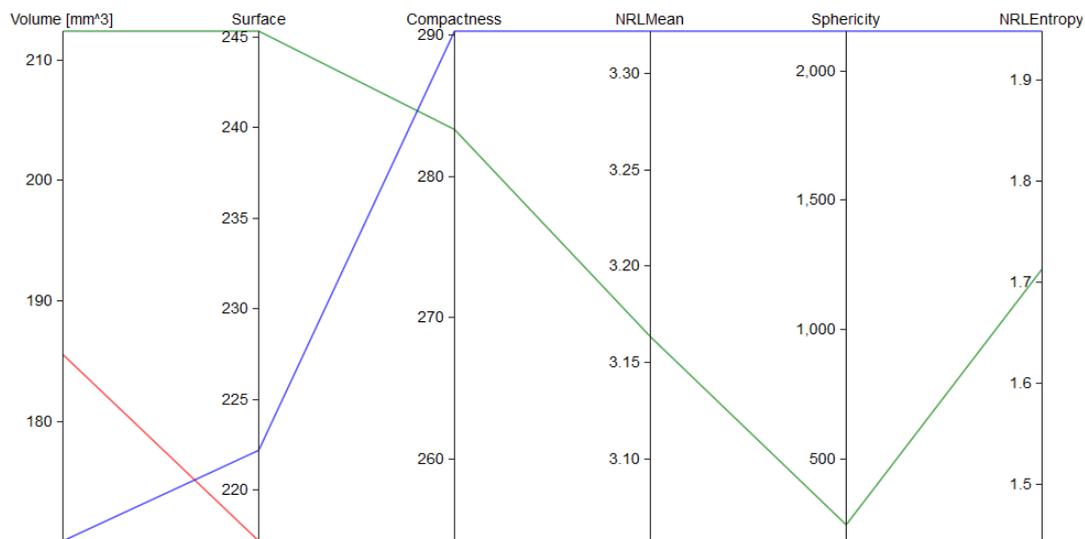


Figure 3.12: PCP of a single patient

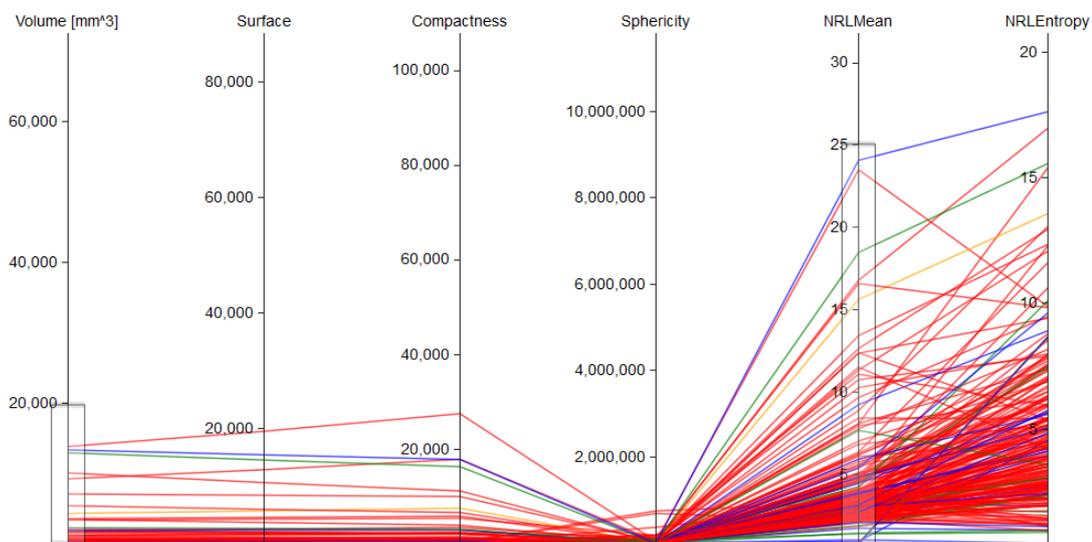


Figure 3.13: PCP of multiple patients with restriction of two features and moved feature axes

Implementation

The implementation of the tool, named CompVis, started in 3DSlicer, but then switched to MeVisLab, Version 3.0 . The reason for the change was the easier to use modular IDE and the ability to connect modules more efficient.

4.1 MeVisLab

MeVisLab is a modular framework for medical image processing focused on scientific visualization. It originated of ILAB developed by Fraunhofer MEVIS, and was transferred to MeVis Medical Solutions AG in 2008.

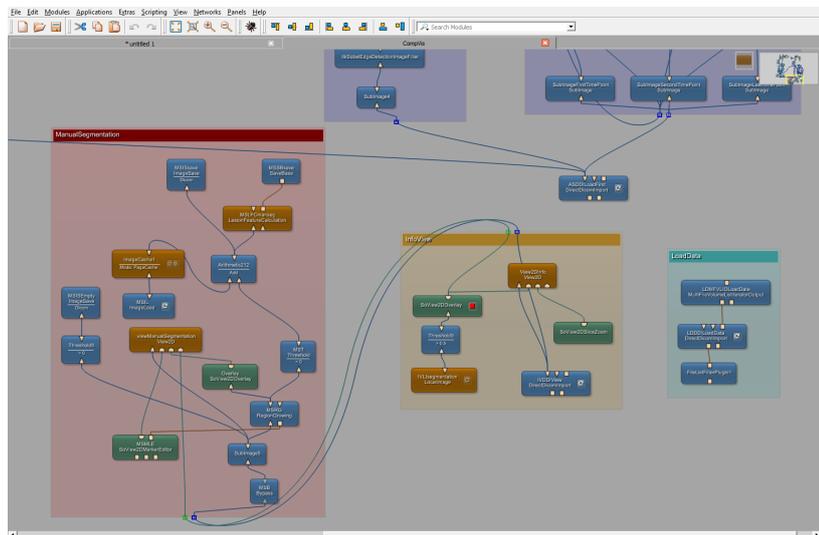


Figure 4.1: MeVisLab IDE example

MeVisLab is written in C++ and it uses the Qt framework for the graphical user interface, its available for Window, Linux and Mac OS X [MeVa].

4.2 Structure and Details

MeVisLab offers an IDE that allows visual programming (Figure 4.1). Figure 4.3 gives an overview of the graphical user interface. On the top are the options to load a study and include additionally data via the csv-format. The loaded study and patients as well as the imported data are then observable on the left table. The right-side features three different views: Segmentation, Intra-patient View and Multi-patient View. The Segmentation tab's purpose is to make the image data available and allow either a manual segmentation by using region growing, combining segmented lesions of a patient by setting an intensity value, or an automatic segmentation using the introduced pipeline of chapter 3.1. Also, the possibility of segmenting multiple patients at once is given by selecting the patients on the left table. The segmentation for each time point of a patient's DCE-MRI data is saved as image and the calculated morphological features as well as the intensity for connection reasons are saved as StandardItemModel exported to an xml file. Therefore, they can be easily reused for other purposes.

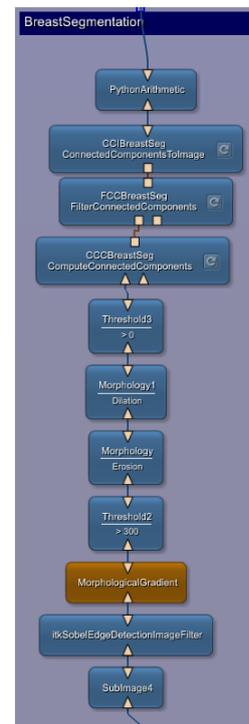


Figure 4.2: Modular construction of the breast segmentation approach in MeVisLab

The Intra-patient view and the Multi-patient view are both based on JavaScript for easy extension. Both views make use of JavaScript libraries Chart.js [Cha] and D3.js [D3j] to implement charts. The Intra-patient view features a calendar view and a bar plot for easy identification of changes of the calculated normalized morphological features. A doughnut chart, to display the diversity and a Sankey diagram to show coherence between different imported data as well as a PCP to visualize selected patients morphological features.

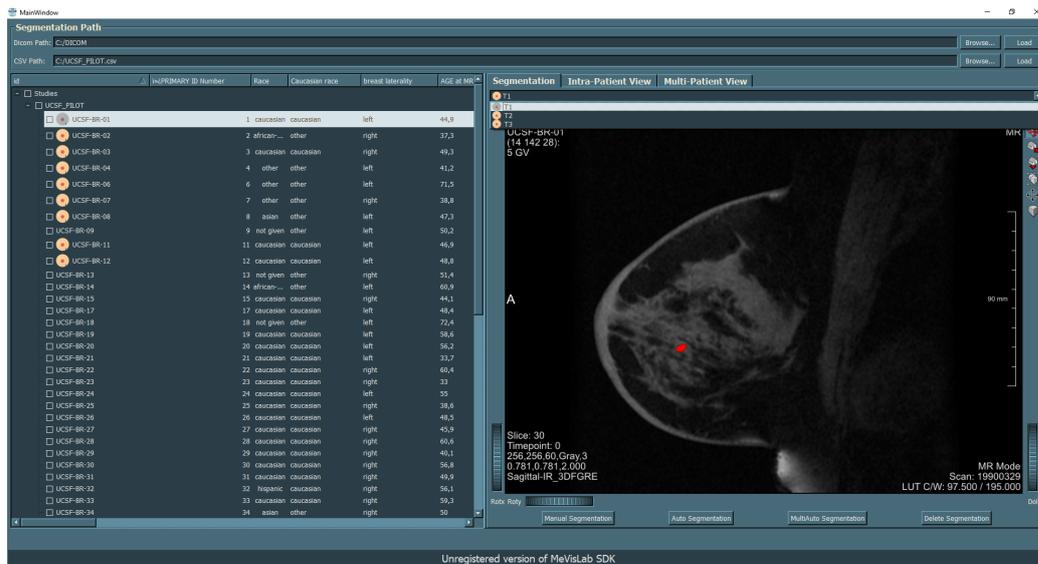


Figure 4.3: CompVis, patient view

4.3 Limitations

The use of python, as main programming language in MeVisLab, lowers the performance compared to an implementation with C++. Also, the missing ability of threading increases the waiting time, especially for loading data and the automatic segmentation process, additionally blocking the graphical user interface resulting in an restricted user experience.

Currently, version 3.0 of MeVisLab has OpenCV Version 3.0 implemented. This version contains a bug, not allowing the SVM to load a trained state. Therefore, the classifier has to be trained at the beginning of each start of the program [Git15].

Due to missing ground truth or manual segmentation, carried out by an expert, the evaluation of the automatic segmentation pipeline was not possible. However, it can be argued, that existing classification systems using only morphological features can lead to highly sensitive, highly specific results [FSF⁺16, TNB03, ASR04].

Conclusion and future work

The following and final section of this thesis discusses the developed methods for lesion segmentation and critically reflects over the elaborated procedures. It summarizes the work of the thesis and gives an overview of future improvements and tasks.

5.1 Conclusion and future work

This work's purpose is to develop a software to support researchers to value applied therapies and to discover new correlations between medical and patient specific parameters. Therefore, a framework was developed to inspect single patients over time and to compare multiple patients' reaction. An automatic segmentation method, which also links the segmented lesions over time, is provided to reduce the extensive work of manual lesion segmentation. For intra-patient and multi-patient analysis an extensible data visualization approach is implemented.

Automatic breast lesion segmentation is a common problem with many different approaches trying to increase sensitivity and specificity of the segmentation. One main issue for these systems is the fact that there is no standardization of the MRI exam. Spatial and temporal resolutions, field of view, image pre- and post-processing, all these are just some examples of variables making it difficult to compare multiple studies. Even though, new papers are able to segment lesions with MR data from different scanners and various parameters [VGD⁺11] a perfect system does not exist at the moment. Automatic lesion mapping over time is a hardly tackled problem. The introduced approach could lead to detecting other lesions or no detect because of the affine transformation. Another problem for researchers is low amount of accessibly DCE-MRI studies [FSF⁺16].

List of Figures

1.1	Cancer statistic, Austria 2015. Data from [Bru18]	1
1.2	Schematic depiction of acquisition procedure. [TCI17]	3
1.3	Example of the scanned DCE-MRI data of a patient	4
3.1	Pipeline of the automatic segmentation mechanism for an individual patient	8
3.2	Comparison of TIC and vesselness filtered images with (left) and without(right) breast segmentation	8
3.3	Breast segmentation pipeline	9
3.4	Visualization of the main phases of the TIC [CL13]	9
3.5	Image after morphological filters (left) and fully filtered image (right). . .	10
3.6	Linear, two-dimensional SVM.	11
3.7	Lesion mapping to subsequent time	12
3.8	Calendar view	13
3.9	Bar plot	13
3.10	Circular Sankey diagram	15
3.11	Use case example of the interactive Sankey diagram	15
3.12	PCP of a single patient	16
3.13	PCP of multiple patients with restriction of two features and moved feature axes	16
4.1	MeVisLab IDE example	17
4.2	Modular construction of the breast segmentation approach in MeVisLab .	18
4.3	CompVis, patient view	19

Bibliography

- [ANN⁺09] Francois Andry, Goutham Naval, Daren Nicholson, Michelle Lee, Igor Kosoy, and Liliya Puzankov. Data visualization in a personal health record using rich internet application graphic components. In *HEALTHINF 2009 - Proceedings of the 2nd International Conference on Health Informatics*, pages 111–116, 2009.
- [ASR04] Lina Arbach, Alan Stolpen, and J.M. Reinhardt. Classification of breast MRI lesions using a backpropagation neural network (BNN). *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano (IEEE Cat No. 04EX821)*, 2:253–256, 2004.
- [Bru18] Brust, 2018.
- [CFR⁺13] Elizabeth Curmi, Richard Fenner, Keith Richards, Julian M. Allwood, Bojana Bajželj, and Grant M. Kopec. Visualising a Stochastic Model of Californian Water Resources Using Sankey Diagrams. *Water Resources Management*, 27(8):3035–3050, 2013.
- [Cha] Chart.js | Open source HTML5 Charts for your website.
- [CL13] Liuquan Cheng and Xiru Li. Breast magnetic resonance imaging: kinetic curve assessment. *Gland surgery*, 2(1):50–3, feb 2013.
- [D3j] D3.js - Data-Driven Documents.
- [EMJ⁺09] Riham H. El Khouli, Katarzyna J. Macura, Michael A. Jacobs, Tarek H. Khalil, Ihab R. Kamel, Andrew Dwyer, and David A. Bluemke. Dynamic contrast-enhanced MRI of the breast: Quantitative method for kinetic curve type assessment. *American Journal of Roentgenology*, 193(4):W295–300, oct 2009.
- [FNVV98] Alejandro F. Frangi, Wiro J. Niessen, Koen L. Vincken, and Max A. Viergever. Multiscale vessel enhancement filtering. pages 130–137, 1998.
- [FSF⁺16] Roberta Fusco, Mario Sansone, Salvatore Filice, Guglielmo Carone, Daniela Maria Amato, Carlo Sansone, and Antonella Petrillo. Pattern

Recognition Approaches for Breast Cancer DCE-MRI Classification: A Systematic Review. *Journal of medical and biological engineering*, 36(4):449–459, 2016.

- [Git15] Github. OpenCV 3.0 Bug, 2015.
- [GMKM⁺15] Albert Gubern-Mérida, Michiel Kallenberg, Ritse M. Mann, Robert Martí, and Nico Karssemeijer. Breast segmentation and density estimation in breast MRI: A fully automatic framework. *IEEE Journal of Biomedical and Health Informatics*, 19(1):349–357, 2015.
- [GSO⁺09] S Glaßer, S Schäfer, S Oeltze, U Preim, KD Tönnies, and B Preim. A visual analytics approach to diagnosis of breast DCE-MRI data, 2009.
- [Ins85] Alfred Inselberg. The plane with parallel coordinates. *The Visual Computer*, 1(4):69–91, aug 1985.
- [JCJG14] Jagadaeesan Jayender, Sona Chikarmane, Ferenc A Jolesz, and Eva Gombos. Automatic segmentation of invasive breast carcinomas from dynamic contrast-enhanced MRI using time series analysis. *Journal of magnetic resonance imaging : JMRI*, 40(2):467–75, aug 2014.
- [JSD⁺10] Jaber Juntu, Jan Sijbers, Steve De Backer, Jeny Rajan, and Dirk Van Dyck. Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. *Journal of Magnetic Resonance Imaging*, 31(3):680–689, mar 2010.
- [Kre17] Brust (C50) - Krebsmortalität (Sterbefälle pro Jahr), Österreich ab 1983. Technical report, Statistik Austria, 2017.
- [LAA⁺14] Xia Li, Lori R Arlinghaus, Gregory D Ayers, A Bapsi Chakravarthy, Richard G Abramson, Vandana G Abramson, Nkiruka Atuegwu, Jaime Farley, Ingrid A Mayer, Mark C Kelley, Ingrid M Meszoely, Julie Means-Powell, Ana M Grau, Melinda Sanders, Sandeep R Bhave, and Thomas E Yankeelov. DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: pilot study findings. *Magnetic resonance in medicine*, 71(4):1592–602, apr 2014.
- [LTR⁺08] Claudette E. Loo, H. Jelle Teertstra, Sjoerd Rodenhuis, Marc J. Van De Vijver, Juliane Hannemann, Saar H. Muller, Marie Jeanne Vrancken Peeters, and Kenneth G.A. Gilhuijs. Dynamic contrast-enhanced MRI for prediction of breast cancer response to neoadjuvant chemotherapy: Initial results. *American Journal of Roentgenology*, 191(5):1331–1338, 2008.
- [MeVa] MeVisLab: History.

- [MeVb] MeVisLab: MeVisLab.
- [NCC⁺08] Ke Nie, Jeon-Hor Chen, Siwa Chan, Man-Kwun I Chau, Hon J Yu, Shadfar Bahri, Tiffany Tseng, Orhan Nalcioglu, and Min-Ying Su. Development of a quantitative method for analysis of breast density based on three-dimensional breast MRI. *Medical physics*, 35(12):5253–62, dec 2008.
- [POM⁺09] Bernhard Preim, Steffen Oeltze, Matej Mlejnek, Eduard Gröller, Anja Hennemuth, and Sarah Behrens. Survey of the visual exploration and analysis of perfusion data. In *IEEE Transactions on Visualization and Computer Graphics*, volume 15, pages 205–220, mar 2009.
- [RHF] Patrick Riehmman, Manfred Hanfler, and Bernd Froehlich. Interactive Sankey Diagrams.
- [RSB92] Jean-Francois Rivest, Pierre Soille, and Serge Beucher. Morphological gradients. page 139, 1992.
- [Sch08] Mario Schmidt. The Sankey diagram in energy and material flow management: Part I: History. *Journal of Industrial Ecology*, 12(1):82–94, feb 2008.
- [SF68] Irwin Sobel and G Feldman. An Isotropic 3x3 Image Gradient Operator for Image Processing. *Machine Vision for Three – Dimensional Scenes*, (June):376–379, 1968.
- [SG14] Marc Streit and Nils Gehlenborg. Points of View: Bar charts and box plots. *Nature Methods*, 11(2):117, 2014.
- [TCI17] TCIA. Breast-MRI-NACT-Pilot, 2017.
- [TCS⁺18] Sneha Thakran, Subhajit Chatterjee, Meenakshi Singhal, Rakesh Kumar Gupta, and Anup Singh. Automatic outer and inner breast tissue segmentation using multi-parametric MRI images of breast tumor patients. *PLoS ONE*, 13(1):e0190348, jan 2018.
- [The81] Martin Theus. High Dimensional Data Visualization. pages 1–7, 1981.
- [TNB03] Angelina A. Tzacheva, Kayvan Najarian, and John P. Brockway. Breast cancer detection in gadolinium-enhanced mr images by static region descriptors and neural networks. *Journal of Magnetic Resonance Imaging*, 17(3):337–342, mar 2003.
- [VGB⁺09] Anna Vignati, Valentina Giannini, Alberto Bert, Massimo Deluca, Lia Morra, Diego Persano, Laura Martincich, and Daniele Regge. A fully automatic lesion detection method for DCE-MRI fat-suppressed breast images. *SPIE medical imaging*, 7260(July):726026–726026–12, 2009.

- [VGD⁺11] Anna Vignati, Valentina Giannini, Massimo De Luca, Lia Morra, Diego Persano, Luca A. Carbonaro, Ilaria Bertotto, Laura Martincich, Daniele Regge, Alberto Bert, and Francesco Sardanelli. Performance of a fully automatic lesion detection system for breast DCE-MRI. *Journal of Magnetic Resonance Imaging*, 34(6):1341–1351, 2011.
- [VV] J.J. Van Wijk and E.R. Van Selow. Cluster and calendar based visualization of time series data. *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99)*, pages 4–9,.
- [WGOM15] Hongbo Wu, Cristina Gallego-Ortiz, and Anne Martel. Deep Artificial Neural Network Approach to Automated Lesion Segmentation in Breast DCE-MRI, 2015.
- [WPI⁺] Lei Wang, Bram Platel, Tatyana Ivanovskaya, Markus Harz, and Horst K Hahn. FULLY AUTOMATIC BREAST SEGMENTATION IN 3D BREAST MRI.
- [WWC⁺13] Shandong Wu, Susan P Weinstein, Emily F Conant, Mitchell D Schnall, and Despina Kontos. Automated chest wall line detection for whole-breast segmentation in sagittal breast MR images. 2013.
- [Yu09] Hwanjo Yu. *Support Vector Machine*, pages 2890–2892. Springer US, Boston, MA, 2009.
- [ZBE⁺07] Yuanjie Zheng, Sajjad Baloch, Sarah Englander, Mitchell D Schnall, and Dinggang Shen. Segmentation and classification of breast tumor using dynamic contrast-enhanced MR images. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 10(Pt 2):393–401, 2007.